

METHOD OF FORMING A THIN FILM USING ATOMIC LAYER DEPOSITION

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to the field of semiconductor devices and, more particularly, to a method of forming a thin film of a semiconductor device using atomic layer deposition (ALD).

2. Description of the Related Art

Forming thin films in state-of-the-art highly integrated semiconductor devices requires many rigorous manufacturing requirements such as a low thermal budget, excellent step coverage, accurate control of film thickness, simple process variables, and low particulate contamination.

Conventional CVD-based methods such as low-pressure chemical vapor deposition (LPCVD), plasma-enhanced chemical vapor deposition (PECVD) are no longer adequate for forming thin films in state-of-the-art devices, meeting the manufacturing requirements. For example, in a typical CVD method, a thin film is deposited at a relatively high temperature. This is undesirable because of the possibility of adverse thermal effects on the devices. Also, the CVD thin film often has drawbacks such as a non-uniform thickness, i.e., thickness variations across the surface of the device or particulate contamination.

As for LPCVD, the hydrogen content of a LPCVD thin film is usually high, and step coverage thereof is often unacceptable.

The atomic layer deposition (ALD) process has been proposed as an alternative to such conventional thin film formation technologies because the ALD process can be performed at lower temperatures than conventional CVD-based methods and also exhibits excellent step coverage.

One such ALD process technology is disclosed in U.S. Pat. No. 6,124,158. Here, a first reactant is introduced to react with the treated surface to form a bonded monolayer of reactive species. A second reactant is introduced to react with the surface to form a desired thin film. After each step in the cycle, the reaction chamber is purged with an inert gas to prevent reaction except on the surface. Typically, the

supplying of the reactant and the purging are implemented under the same pressure due to reasons such as the maintenance of fabrication equipment.

However, such conventional ALD technologies also have several drawbacks such as low throughput due to problems such as a relatively low growth rate of atomic layers. Further, the reaction space of conventional ALD reactors such as a traveling wave-type reactor is designed to be very small to reduce the purging volume for purging of byproducts or the like. Thus, conventional ALD reactors only process one or two wafers in each operation, typically one substrate for one operation in a single reactor. Such drawbacks have made it difficult for many conventional ALD technologies to be put in practical applications and commercially acceptable, i.e., mass production.

Recently, several attempts have been made to increase the throughput of the ALD process. One such attempt is disclosed in U.S. Pat. No. 6,042,652. Here, the ALD reactor includes a plurality of modules and a plurality of reaction spaces (stages), i.e., spaces partitioned with the plurality of assembled modules. For example, a lower module is placed below an upper module, thereby creating one reaction space (one stage) between them, which is capable of receiving merely one semiconductor substrate.

However, because each reaction space (stage) is small and partitioned, i.e. being separated from each other, each substrate is inserted in one of the reaction spaces (stages) one by one. Thus, it is hard to utilize an automated wafer transfer mechanism for loading/unloading of the plurality of wafers. Consequently, it takes a significantly long time to load/unload wafers. Also, the number of wafers that can be loaded and processed is still not sufficient.

Accordingly, what is clearly needed is a novel ALD process enabling high throughput that can deal with above mentioned problems while still providing high-quality thin films.

SUMMARY OF THE INVENTION

The present invention provides a method of forming a thin film using atomic layer deposition (ALD). A reactor having a single reaction space is provided. A batch of substrates is concurrently loaded into the single reaction space of the reactor.

Then, a gas containing reactants is introduced into the single reaction space, and a portion of the reactants is chemisorbed on top surfaces of the batch of substrates or wafers within the single reaction space. Non-chemically adsorbed reactants are then removed from the single reaction space.

5 In accordance with one embodiment of the present invention, after introducing the gas containing reactants, non-chemically adsorbed reactants are diluted in the single reaction space to facilitate the removal of non-chemically adsorbed reactants.

Also, according to another embodiment of the present invention, a method of forming a thin film is disclosed in which a reactor having a single reaction space is provided. A plurality of wafers, each having a processing surface, is introduced into the reaction space. The processing surfaces of the plurality of wafers face in substantially the same direction. First reactants are introduced into the reaction space such that a portion of the first reactants is chemically adsorbed on the processing surfaces of the plurality of wafers for ALD. Then, a non-chemically adsorbed portion of the first reactants is removed from the reaction space. Next, second reactants are introduced into the reaction space. Also, a portion of the second reactants is chemically adsorbed on the processing surface of each of the plurality of wafers. Subsequently, a non-chemically adsorbed portion of the second reactant is removed from the reaction space.

20 The foregoing and other objects, features and advantages of the invention will become more readily apparent from the following detailed description of a preferred embodiment of the invention that proceeds with reference to the accompanying drawings.

25 BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a schematic cross-sectional view of an ALD reactor in accordance with one embodiment of the present invention.

FIG. 2 is a graph illustrating the pressure of the ALD reactor in each step of ALD in accordance with one embodiment of the present invention.

30 FIGS. 3A-3D illustrate the process steps for forming an ALD thin film according to one embodiment of the present invention.

FIG. 4 is a graph illustrating process conditions in accordance with one embodiment of the present invention.

FIG. 5 is a graph illustrating the result of the ALD process performed in accordance with one embodiment of the present invention.

FIG. 6 is a graph illustrating the result of the ALD process performed in accordance with one embodiment of the present invention.

5

DETAILED DESCRIPTION

The present invention in general contemplates a method of manufacturing a thin film by an ALD technology, through which throughput can be significantly increased compared with the conventional ALD technologies.

In the following description, numerous specific details are set forth to provide a through understanding of the present invention. However, one having ordinary skill in the art should recognize that the invention can be practiced without these specific details. In some instances, well-known process steps, and techniques have not been shown in detail to avoid obscuring the present invention.

Herein below, a method of forming a thin film using ALD in accordance with an embodiment of the present invention is described.

Referring to FIG. 1, there is schematically shown an ALD reactor 10 having a single reaction space 12 within a process tube 11. Other parts of the reactor 10 such as a heater are omitted for the sake of simplicity. Preferably, the ALD reactor 10 is a furnace-type vertical reactor (vertically oriented) that is similar to conventional LPCVD furnaces as shown in U.S. Pat. Nos. 5,217,340 and 5,112,641. However, any other type of reactor such as one being horizontally oriented, which is suitable for implementing the present invention, can be used instead within the spirit and scope of the present invention.

According to the present invention, the reaction space 12 can mean a space where substrates 15 (or wafers) are placed and where various process sequences of ALD occur. Also, in the present invention, the single reaction space 12 is not partitioned or separated. This is different from the reaction space of conventional reactors such as one shown in U.S. Pat. Nos. 6,042,552 and 6,015,590, where there are multiple (partitioned) reaction spaces in an ALD reactor. In these conventional ALD reactors, particularly, in U.S. Pat. No. 6,015,590, because each of the multiple (partitioned) reaction spaces has a very narrow cross section for minimizing the volume of the reaction space for purging efficiency, the number of the substrates that

can be placed in each of the reaction space is very small, for example, one or two substrates for each reaction space. Also, this aspect of conventional ALD reactors limits the total number of substrates that can be placed in a reactor due to such structural limitations as described above. For example, the modules themselves
5 shown in U.S. Pat. No. 6,042,652 which form each reaction space can take up a significant amount of space or volume within the reactor. These can severely reduce the throughput of the ALD process.

However, in the present invention, because the furnace-type ALD reactor 10
10 has a large-volume single reaction space 12 without being partitioned, the ALD reactor 10 can receive more than one hundred (100) substrates therein as illustrated in FIG. 1. Thus, the number of the substrates to be processed for one ALD operation can be significantly increased (significant increase of throughput).

To process the substrates 15 forming ALD thin films thereon, one batch 14 of
15 substrates 15 is substantially concurrently loaded into the single reaction space 12 of the ALD reactor 10 as schematically illustrated in FIG. 1. In the present invention, the one batch 14 can mean the total number of substrates loaded into the reactor 10 for one ALD operation to form a thin film on the substrates 15. In accordance with one embodiment of the present invention, the one batch 14 preferably comprises approximately 125-135 substrates. Each of the substrates 15 has a processing surface
20 17 preferably on its top.

Preferably, in the present invention ALD process, during the load/unloading of the substrates 15, the batch 14 of substrates 15 is loaded to the ALD reactor 10 using an automatic (i.e., non-manual) wafer transport mechanism 18 as schematically
25 illustrated in FIG. 1. Such automatic wafer transport mechanism 18 can be one disclosed in U.S. Pat. Nos. 5,217,340 and 5,112,641. However, any other type of automatic wafer transport mechanism that is suitable for implementing the present invention can be used instead within the spirit and scope of the present invention.

In other words, in the present invention, because all of the product substrates
30 15 for a single ALD operation can be placed in the single reaction space 12, not distributed among several reaction spaces in a reactor, the loading/unloading of a batch 14 of substrates 15 can be automatically and quickly done by the wafer transport mechanism 18. In particular, the batch 14 of substrates are arranged in a predetermined manner and inserted into a boat 19. The boat 19, typically formed of quartz or other conventional materials, has a plurality of grooves on its inner surface

to receive each of substrates 15. And the boat 19 containing the batch 14 of substrate is loaded into the ALD reactor 10, thus concurrently loading the batch 14 of substrates 15 into the single reaction space 12 of the ALD reactor 10 in a manner schematically described in FIG. 1. Here, substantially all the top surfaces 17 (processing surfaces) of the substrates 15 face the same direction for automated wafer transfer.

This provides an important advantage, especially in terms of throughput, over conventional ALD techniques such as one disclosed in U.S. Pat. 6,015,590, where top surfaces of the wafers faces opposite direction, thus automatic wafer transfer is quite cumbersome or impossible. Thus, in the conventional ALD technologies, only a small number of substrates, mostly one, can be placed in each reaction space one after another. This is because the substrates need to be distributed among several reaction spaces in a reactor and the distribution is nearly impossible or difficult to be done all at once. This is also true for a conventional ALD technique disclosed in U.S. Pat. No. 6,042,652, where a plurality of circular semiconductor substrates are transferred one by one into the reaction spaces (stages) as described above in the background. This whole loading process takes a long period of time, significantly reducing the throughput and thus limiting the commercial application of the ALD process.

As schematically illustrated in FIG. 3A, as in the manner of conventional ALD technologies, first reactants 40 or a gas containing first reactants is introduced (dosing step) through the input 16 of FIG. 1 such as gas supply lines (not shown) of the ALD reactor 10 into the single reaction space 12. Thus, a portion of the first reactants 40 is chemically adsorbed (chemisorbed) onto the processing surfaces 17 of the batch 14 of the substrates 15 within the single reaction space 12. As shown in FIG. 2, the dosing step 31 is preferably performed at a first predetermined pressure P₁, between approximately 0.1 Torr and approximately 0.5 Torr.

On the other hand, in the present invention, to further increase the throughput of ALD, the purging time for ALD need to be reduced. This is because, in general, the purging time is dependent on the volume of a reactor. Because the present invention utilizes a furnace-type reactor having a large volume, the purging volume is substantially larger than other conventional ALD techniques such as the traveling wave type apparatus shown in U.S. Pat. Nos. 6,042,552 or 6,015,590.

To deal with this issue, according to one embodiment of the present invention, after introducing the first reactants 40, to effectively reduce the purging time, a non-

chemically adsorbed portion of the first reactants 40 is diluted in the single reaction space 12 before removal of the non-chemically adsorbed portion of the first reactants 40 from the ALD reactor 10. Here, the non-chemically adsorbed portion of the first reactants 40 includes physically adsorbed (physisorbed) reactants, i.e. first reactants 40 physically attaches onto, and is loosely held to the chemisorbed portion of the first reactants 40 or any residual reactive material within the ALD reactor 10.

For the diluting step 33 of FIG. 2, the ALD reactor 10, as shown in FIG. 1, includes a pressure control valve 21 connected to an exhaustion line 25 or a roughing line for removing the diluted non-chemically adsorbed portion of the first reactants 40 from the ALD reactor 10. The exhaustion line 25 is connected to a pump 23 for discharging the non-chemically adsorbed portion of the first reactants 40 from the reactor 10 to the outside. During the diluting step 33, the control valve 21 is substantially closed and an inert gas is supplied through input 16 into the reactor 10, and the introduction of the first reactants 40 into the ALD reactor 10 is substantially stopped. That is, conductance of the exhaustion line 25 of the ALD reactor 10 is reduced.

Alternatively, during the diluting step 33, an inert gas with an amount substantially more than the amount of the first reactants 40 is introduced into the ALD reactor 10 while stopping the introduction of the first reactants 40 into the reactor 10.

Preferably, as illustrated in FIG. 2, during the dilution of the non-chemically adsorbed portion of the first reactants 40, the reactor pressure is increased from the first predetermined pressure P1 to a second predetermined pressure P2 and thus the second predetermined pressure P2 is greater than the first predetermined pressure P1. Preferably, the second predetermined pressure P2 is greater than approximately 1.5 times the first predetermined pressure P1.

These steps allows the non-chemically adsorbed portion of the first reactants 40 in the reactor 10 to be diluted in a very short period of time, for example, a few seconds, thus drastically reducing the overall purging time and purging efficiency during a purging step 32 compared to conventional ALD techniques. This dilution process significantly reduces the partial pressure of the non-chemically adsorbed portion of the first reactants 40 in the ALD reactor 10. Thus, only a very small amount of the non-chemically adsorbed portion of the first reactants 40 remain in the reactor 10 after the removal of the non-chemically adsorbed portion of the first reactants 40 as the reactant 40 is already diluted, thus maximizing purging efficiency.

Also, because the first reactants 40 are diluted, intermixing between the first reactants 40 can be sufficiently prevented.

Then, as illustrated in FIG. 3B, the diluted non-chemically adsorbed portion of the first reactants 40 is removed (evacuated) from the single reaction space 12 before introducing second reactants 42 (dosing step 35) to form a desired ALD thin film 44 of FIG. 3D by chemical exchange. Preferably, removal of non-chemically adsorbed portion of the first reactants 40 is performed by pumping the reactor 10 using the pump 23, thereby lowering the pressure of the reactor 10 to a third predetermined pressure P3 (See FIG. 2). The third predetermined pressure P3 is lower than the first predetermined pressure P1 of the dosing step 31. Preferably, the third predetermined pressure P3 is lower than approximately 0.5 times the first predetermined pressure P1.

During this step, the decrease of the pressure to the third predetermined pressure P3 can be accomplished by stopping or reducing the introduction of the inert gas and opening the control valve 21. That is, conductance of the exhaustion line is increased.

Now referring to FIG. 3C, the second reactants 42 are introduced into the reaction space 12, and thus a portion of the second reactants 42 is chemisorbed on the processing surfaces 17 of the batch 14 of the substrates 15 to effect chemical exchange. Of course, the diluting step 37 is preferably performed after the dosing 35 of the second reactants 42.

Turning to FIG. 3D, a non-chemically adsorbed portion of the second reactants 42 is removed (removing step 38 of FIG. 2) from the reaction space 12 using the same method as described above that is applied to the first reactants 40 during the removing step 34.

The above steps of introduction of the first and second reactants 40, 42 and removal of non-chemisorbed portions of the reactants 40, 42 from the reaction space 12 can be repeated to achieve a desired film thickness.

It is to be noted that the purging method suggested by the present invention is not dependent on the kind of the reactant, and thus can be utilized for formation of various ALD thin films. Such ALD thin films are, for example, an oxide layer of Al_2O_3 , TiO_2 , ZrO_2 , HfO_2 , Ta_2O_5 , Nb_2O_5 , CeO_2 , Y_2O_3 , SiO_2 , In_2O_3 , RuO_2 , or IrO_2 . The other examples are as follows: a composite oxide layer of SrTiO_3 , PbTiO_3 , SrRuO_3 , CaRuO_3 , $(\text{Ba},\text{Sr})\text{TiO}_3$, $\text{Pb}(\text{Zr},\text{Ti})\text{O}_3$, $(\text{Pb},\text{La})(\text{Zr},\text{Ti})\text{O}_3$, $(\text{Sr},\text{Ca})\text{RuO}_3$, $(\text{Ba},\text{Sr})\text{RuO}_3$, Sn doped In_2O_3 (ITO), Fe doped In_2O_3 , or Zr doped In_2O_3 ; a nitride

layer of SiN, NbN, ZrN, TiN, TaN, Y_3N_5 , AlN, GaN, WN, or BN: a complex nitride layer of WBN, WSiN, TiSiN, TaSiN, or AlTiN: a metal layer of Si, Al, Cu, Ti, Ta, Mo, Pt, Ru, Rh, Ir, W, or Ag: a silicide layer of Al, W, Ti, or Co: and a metal silicate material ($M_{1-x}Si_xO_2$). Here, the metal "M" can be hafnium (Hf), zirconium (Zr), tantalum (Ta), titanium (Ti), Cesium (Cs) or aluminum (Al). Those of skill in the art will appreciate that lists are not exhaustive or exclusive and are not intended to limit in any way the scope of the invention as claimed.

Example 1

A SiN film is deposited by the present invention ALD process. The reactants used are DCS ($SiCl_2H_2$) and NH_3 gases activated by a remote plasma (400W). The deposition temperature is $375^\circ C$. Flow rates of the reactants are 500 sccm for DCS and 2000sccm for NH_3 . For dilution before removal of non-chemically adsorbed reactants, 5000 sccm of N_2 gas is introduced into a reactor. The time and pressure for each step of DCS supplying, DCS purging, NH_3 supplying and NH_3 purging is shown in table 1 and is further illustrated in FIG. 4. Also, FIG. 5 illustrates the results of the above ALD process.

Table 1

Item	DCS dosing		DCS purging		NH3 dosing		NH3 purging	
			Diluting	Removing			Diluting	Removing
Pressure (Torr)	$<0.1 \rightarrow 2.0$	2.0	$2.0 \rightarrow >10$	$>10 \rightarrow <0.1$	$<0.1 \rightarrow 0.25$	0.25	$0.25 \rightarrow >10$	$>10 \rightarrow <0.1$
Time (sec)	7.5	10	4	6	1.5	10	4	6

The growth rate in accordance with the ALD process described above is 1\AA /cycle, and a good ALD processing characteristic can be obtained.

Further, it has been observed that, without using the purging method of the present invention, the following problems could occur. First, if the purging step is performed with an inert gas such as Ar or N₂, under the same pressure as the pressure during the dosing step of the reactants, a substantial amount of the inert gas can remain in the reactor. This reduces the partial pressure of the reactants. Accordingly, the reactant dosing time for the next dosing step can be increased. In addition, the purging time is also increased. Second, if the purging is performed by pumping only without dilution before the pumping as in the one embodiment of the present invention, the purging takes significantly a long time.

Example 2

HCD (Si₂Cl₆) is stored in a bubbler at room temperature and is introduced to a reactor with 500 sccm of N₂ gas as a carrier gas. Then, a purging is performed by diluting non-chemically adsorbed reactants with 5000 sccm of N₂ gas and then pumping (removing) away the non-chemically adsorbed reactants from the reactor. Next, 2000 sccm of remote plasma (400W) NH₃ is supplied and the purging is done by diluting non-chemically adsorbed reactants with 5000 sccm of N₂ gas and then pumping (removing) away the non-chemically adsorbed reactants from the reactor.

At this time, HCD is supplied to the reactor for 20 seconds. The reactor pressure is changed from 0.1 torr to 2 torr and then kept at 2 torr. The pressure during the purging was changed from 2 torr to 10 torr during the dilution step (4 seconds) and then, lowered to 0.1 torr during the pumping (6 seconds). The Supplying of NH₃ (30 seconds) and purging (4 + 6 seconds) was implemented by the same manner as described above. FIG. 6 illustrates the results of the above ALD process.

The growth rate was 2.3 Å/cycle and a good ALD processing characteristic can be obtained.

Some of the characteristics of the present invention described above are as follows:

1. The reactor pressure during the dosing step and that during the purging step can be different.

2. The reactor pressure for each dosing step for different reactants can be substantially the same or different.

3. The purging step may be seen to include a diluting step where the reactor pressure increases from the pressure during the dosing step of the reactants, and a removing or evacuation step of which pressure decreases to a lower pressure than the pressure during the dosing step of the reactants.

5 By utilizing these characteristics, the following effect can be obtained.

1. The dosing step of each reactant is dependent of partial pressure and time (reactant exposure dependency as illustrated by Langmuire). Accordingly, by increasing the partial pressure of the supplied reactant during the reactant dosing, the processing time can be reduced.

10 2. Differently from the conventional ALD process where a constant pressure should be kept, the dosing step of each reactant is accomplished after the purging by the pumping. Therefore, a desired pressure can be obtained from the lower pressure.

15 3. When the purging is implemented within a reactor having a large volume, an inert gas is firstly applied to dilute the reactants. Then, the pumping is implemented to obtain the desired degree of the purging effect within a short period of time.

In conclusion, the present invention offers many advantages over the conventional ALD techniques and overcomes many drawbacks of the conventional ALD techniques. For example, the present invention significantly increases the throughput of the ALD process. In particular, according to preferred embodiments of the present invention, because the furnace-type ALD reactor of the present invention has a large-volume single reaction space without being partitioned, it can receive and process more than 100 substrates all at once, substantially more than any other conventional ALD techniques. Also, because all of the product wafers for one ALD process can be placed in a single reaction space, not distributed among several reaction spaces in a reactor, the loading/unloading of the batch of substrates can be automatically and quickly done by an automatic wafer transport mechanism. Further, non-chemically adsorbed reactants are diluted in a single reaction space before removal of non-chemically adsorbed reactants from the reaction space, the purging time can be significantly reduced and purging efficiency is maximized.

Besides these advantages, the present invention ALD reactor is less costly than conventional ALD reactors and is easier to maintain. Thus, the present invention ALD process increases throughput and manufacturability to the extent that mass production is possible with ALD.

